DOCUMENT RESUME

ED 085 618 CG 008 524

AUTHOR Sarri, Rosemary C.; Selo, Elaine
TITLE Evaluation Process and Outcome in Juvenile Corrections: Musings on a Grim Tale.
PUB DATE Mar 73
NOTE 55p.; Paper presented at the Fifth Banff International Conference on Behavior Modification, Banff, Alberta, Canada, March 25-29, 1973

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Corrective Institutions; Delinquency; *Delinquent Rehabilitation; Evaluation Methods; Literature Reviews; *Performance Criteria; *Program Effectiveness; *Program Evaluation; Research Problems

ABSTRACT

This paper is a review of some of the issues, dilemmas, and constraints in the evaluation of juvenile corrections. Included are such problems as: (1) adequate assessment criteria; (2) insufficient controls in experimental designs; (3) the conflict between humaneness and effectiveness; and (4) the measurement of program goals. A critique of six studies in juvenile correctns provides examples whose strengths and weaknesses can serve as quicklines for further refinement and elaboration of evaluation methodologies. The process components of these studies, rather than their outcomes, are emphasized. The thoroughness with which the researchers evaluated programs is summarized in a series of charts, one for each of five dimensions; (1) subject population; (2) setting; (3) treatment technology; (4) process variables; and (5) outcome measures. The authors conclude with a delineation of a plan for comparative assessment involving the classification of the functional categories of juvenile corrections into four major groups: (1) detention programs; (2) processing; (3) change and control; and (4) re-integration programs. (RWP)

FILMED FROM BEST AVAILABLE COPY

ED 085618

The University of Michigan

National Assessment of Juvenile Corrections

EVALUATION PROCESS AND OUTCOME IN JUVENILE CORRECTIONS: MUSINGS ON A GRIM TALE

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Rosemary C. Sarri
Elaine Selo

CG008524

Paper prepared for the Fifth Banff International Conference on Behavior Modification, Banff, Alberta, Canada, March 25-29, 1973.

> Once upon a time, 1940 to be exact, a sociologist wrote of a strange community to which men were temporarily banished because they were possessed by evil spirits. Although high priests were sent to visit the banished men, their efforts to drive out the evil spirits proved in vain. The banished men resisted the efforts of the high priests and withdrew from them, speaking in a strange language and living by rules foreign to the high priests. Under these conditions, the evil spirits in many of the men, instead of withering away, increased and multiplied. Thus, when the men were finally allowed to return to the land from which they came, the people found them possessed by spirits more numerous and more evil than before, and they caused the men to be banished again and again. (Slosar, 1972)

More than thirty years ago, Clemmer (1940) presented his prisonization hypothesis about the effects on human beings from isolation in closed institutions. Although the prison epitomizes the extreme in corrections programs, it is probable that all correctional programs are afflicted with the "grim reality" of ineffectiveness in the rehabilitation of offenders. Certainly juvenile corrections is no exception! Evaluation of program processes and outcomes has been less extensive than in adult corrections, but it is still apparent that juvenile intervention seldom succeeds to anyone's satisfaction. The juvenile justice system is falling far short of its objectives: serving the best interests of individual youth and contributing to public safety by controlling and reducing youthful crime. Disappointment with juvenile justice is especially strong because of the humanitarian hopes generated by the founding of the juvenile court at the beginning of the century. One method for resolving at least part of the difficulty is to focus greater resources on thorough evaluation of correctional programs so that decisions can be made on the basis of greater knowledge about the probability of attaining a given outcome from a specified program of intervention. As in other human service organizations today, evaluation is one of the "in" activities in juvenile corrections. Legislators and boards increasingly request that there be systematic evaluation of newly-funded programs. Unfortunately, very little of the activity that is subsumed under "evaluation" could be classified as research, but there is a consistent groping for more effective methods of intervention. Any review of the evaluative

---

Although the authors are responsible for the content of this paper, they are especially appreciative of the comments and criticisms of their colleagues in the National Assessment of Juvenile Corrections, especially to Y. Hasenfeld, A. McNeece, W. Gricht and R. D. Vinter.

research literature in juvenile corrections would draw essentially the same conclusions as those drawn by Rossi in his assessment of the evaluation of poverty, education, and other social action programs (Rossi and Williams, 1972). He concluded that there were few sufficiently powerful research designs among those studies to permit unequivocal statements that could be used for policy formulation.

This paper reviews some of the issues, dilemmas, and constraints in the evaluation of juvenile corrections. It examines the implications of organizational goals for the evaluation of processes and outcomes, and following that, a series of contracting studies of juvenile corrections are analyzed with reference to their goals, characteristics of subjects, treatment technologies, organizational effort and process, and outcomes. Societal values as a constraint on criteria for assessment and on means of intervention are considered along with particular problems of measurement in this category of human service organizations. Lastly, elements of the plan are proposed for the evaluation in which we are engaged at the present time. When this plan is fully operationalized, we hope that it will enable us to assess significant aspects of the effectiveness of variant types of juvenile correctional programs in a large number of states.

ISSUES AND DILEMMAS IN EVALUATION

What should be evaluated in the assessment of juvenile corrections? At least three orders of phenomena are of importance in any evaluation that has explanatory, as well as policy, implications:

(1) the personal and social characteristics of the target population that the organization seeks to change;

(2) the structures and practices within the organization that must be implemented if an offender is to be changed from Condition A to Condition B;

(3) the inter-organizational exchange among units within or linked to the juvenile justice system that have consequences for the varying careers of juvenile offenders.

Most evaluative research to date has addressed the personal and social characteristics of the individual offender -- personality, behavior, values, attitudes and

capabilities (Schrag, 1971; Suchman, 1967). Insufficient attention has been directed to organizational goal implementation, stability and adaptability, technological feasibility, referral rates, and organizational structures required for quality performance (Mott, 1972). Programs typically are judged as effective or ineffective by reference only to individual level results (California Youth Authority, 1973). Moreover, even the latter type of evaluation often fails to consider sufficiently the selective input of offenders into different types of programs. When selective assignment occurs, as it does in most correctional classification, methods must be devised for taking this factor into consideration before there can be comparative evaluation across programs or organizations. "Tracking" is an observable phenomenon in juvenile corrections as it is in public schools and other human service organizations (Cicourel and Kitsuse, 1964; Berghardt, et al, 1971; Sarri et al, 1970). Again, individual characteristics are used as the basis for assignment to a given program which has a greater or lesser probability for successful outcome, independent of those individual characteristics. For example, a juvenile assigned to probation has greater opportunities for education and employment than most youth assigned to institutions. Yet the criteria for assignment may be race, sex, family composition and so forth, characteristics unrelated to ability to take advantage of education or employment. Furthermore, "time of the year" is a factor that may affect the processing of a juvenile. The probability of commitment by a juvenile court to an institution is lower in the spring quartile than in other months of the year because the court is likely to have expended its resources toward the end of the fiscal year and cannot afford institutionalization. In addition to the above factor, there are many other organizational conditions which should be considered in evaluation, but too often they are ignored or an assumption about randomized effects is proposed as a general explanation.

Standards of performance at the organizational and inter-organizational levels have been almost non-existent in corrections; where they were found, they were based largely on the collective subjective opinions of administrators. Objective criteria and on-going program evaluation are only now being developed. At least five categories of criteria can be delineated for organizational evaluation to then be linked with

outcome evaluation. These include: (1) effort (e.g., cost, time, and types of personnel expended to achieve goals); (2) level of performance (e.g., the number of individuals who complete Program X from those who enroll); (3) adequacy of performance (e.g., value of the program to offenders); (4) efficiency (e.g., relative cost of Program X); and (5) organizational processes (e.g., program attributes which relate to success or failure, recipients who benefit, service delivery, and first order effects produced during the period of intervention).

Recidivism: A Criterion for All Seasons

A second and equally important question is: Should recidivism be the primary outcome criteria in the evaluation of juvenile corrections, and if so, how is it to be defined? When one thinks of effectiveness and outcomes in corrections, the concept "recidivism" inevitably is considered. Most often it is used to refer to some absolute measurement of post-program law-violative or morally disapproved behavior. This is particularly problematic in the case of juvenile offenders because data about recidivism reinforces the "criminal" label; yet, a large proportion of juvenile offenders in many programs are guilty only of status offenses -- truancy, running away, promiscuity, incorrigibility, and so forth. None of these behaviors are crimes for adults, but juvenile "recidivists" are labeled as law-violators along with adults who commit felonies. Even in states or communities where status offenses are not sufficient basis for state intervention, their repeated commission may be a basis for the label "recidivist" and for incarceration, often in jails and for longer periods of time than for offenders who do violate the law.

Another problem associated with the use of recidivism in absolute rather than in relative terms is that positive results may be obscured. Thus, a single arrest or violation is enough to classify the person as a "recidivist" and an outcome failure. If recidivism is conceived in relative terms, there is an expectation that the program will result in fewer and less serious offenses by participants. The addition of other outcome measures reduces reliance on recidivism as an absolute criteria. Thus, positive changes in education, employment, family life, and so forth may be linked to program experiences. The recently published results of the Provo and

Silverlake Experiments (Empey and Erickson, 1972; Empey and Lubeck, 1971) indicate very clearly a relative reduction in the seriousness and frequency of law-violative behavior when pre-program results are compared with post-program behavior. Furthermore, they were also able to identify important differences in behavior in relation to age. Their findings are similar to those obtained by Miller (1962) in a study of a community delinquency prevention program. Both observed that the program results varied according to the age of the youth at the time of entrance and exit from the program. These patterns could be linked to general patterns of criminal behavior for all youth at different ages. Thus, youth who entered the program at younger ages would be expected to demonstrate a slower reduction of criminal behavior from those who entered and left at older ages because criminal behavior peaks in the late teens and then tapers off sharply. It is also possible that other general developmental patterns of youth will produce variable outcome patterns from similar program experiences. When known, these can be considered in the program design and in the development of "base expectancy" criteria for the categorization of populations according to the probability of success or failure in the community. Moberg and Ericson (1972) argue that if recidivism is to be used meaningfully as an outcome criteria, it must be conceptualized as a continuum with variable probabilities developed for programs and individuals. They also assert that it is wholly unrealistic to expect "total conversion" as we apparently do at the present time in measuring change from criminal to non-criminal behavior. For purposes of program planning, knowledge about the length of time and phases of change are as important as ultimate outcomes.

Recidivism is also a problematic criterion for the measurement of effectiveness because it obscures differentiation between short-run and long-run consequences for the society. In the case of youth, that intervention which may protect society in the short-run (e.g., incarceration in a closed institution) may have long-run negative consequences and vice versa. The effectiveness of any intervention is to be assessed on the basis of its ability to rehabilitate and reintegrate juveniles into meaningful social roles, not merely to reduce immediate law-violative behavior. Furthermore, over-reliance on absolute measures of recidivism obscures negative or

no impact from correctional experiences. Far too few studies ever measure negative impact other than recidivism, despite general awareness of the inevitability of this outcome in many programs. Were these data to be made as routinely available as is recidivist data, opportunities for change in custodial programs might be vastly increas (Kassebaum, Ward and Wilner, 1970). Only recently in the plethora of legal actions against correctional agencies has there been any substantial effort to measure negative impact.

Ward and Kassebaum (1972) contend that the lack of concern about the overwhelming evidence of negative results from correctional intervention occurs because departments of corrections are concerned more with surveillance and control than rehabilitation. Thus, it makes little sense to measure their present outcomes in terms of rehabilita individuals, because that is not their primary goal. Obviously, much evaluation has failed to address this phenomenon.

The unreliability of crime statistics and the manipulation of data by administrative boards, police and other agencies is another reason for caution in the use of recidivist data in the evaluation of correctional programs. Police, courts, parole boards, and other agencies may seek to improve or depreciate the public status of a particular program and, in order to do so, the data about individual behavior may be covertly manipulated. The unreliability of crime statistics was recently documented in an analysis by Seidman and Couzens (1972). They analyzed police reports in several metropolitan communities and observed reports of reductions in selected types of crime which corresponded with political pressures to reduce such crime. Thus, they conclude "crime statistics...are highly misleading indicators of what they are used to measure, at least in part, simply because they are used as measures" (p.29). It has also been recognized for a long time that persons in programs or releasees may be surveilled and harrassed by law enforcement officials far more than the rest of the population.

Experimental Design Dilemmas

Thus far, we have been concerned with issues in evaluation that involve measurement of outcome. Another set of issues which has been discussed extensively in

recent years in the evaluation literature concerns problems in the use of experimental designs (Weiss, 1972; Rossi and Walker, 1972; Caro, 1971). Seldom do researchers in juvenile corrections have the authority or resources to exert the necessary controls required by powerful experimental designs, nor are they able to measure objectively ultimate and process goals as these relate to behavioral outcome criteria. Many assume that correctional decision-makers can select variable types of programs for youthful offenders; thus, there should be few limits on application of experimental designs in evaluation. Seldom, however, do administrators have such discretion; in fact, in many communities, judges wholly control where juveniles will be placed, for how long, and when they may return to their home communities.

Campbell's (1966) proposals for quasi-experimental designs offer some solutions for sound evaluative research, as do those by Guttentag (1971) for decision-theoretic models. Assumptions about interchangeability of units and definitions of variables must be made with great caution. Guttentag (1971) points to some of the problems which occurred in the Westinghouse Study of Headstart programs where the assumption was made that these programs throughout the country could be viewed as unitary variables. Our field research in juvenile corrections in several states indicates clearly that regional, cultural, and socio-economic differences must be examined and controlled if comparisons are to be made among organizations and persons. Furthermore, statutory variations and judicial decisions may dramatically alter conditions in the middle of an evaluation experiment. It is inevitable that much research in corrections will continue to be done on a non-experimental basis, but knowledge about effectiveness in corrections can evolve as a consequence of multiple types of experimentation rather than from a few definitive classical experiments.

Humaneness and Justice vs. Effectiveness

Recent statements by organizations such as the American Friends Service Committee (1971), the American Association of University Women (1970), by social researchers such as Lerman (1968), as well as judicial decisions have highlighted consideration of the relationship between effectiveness and conditions of fairness,

humaneness, and justice in the operation of correctional programs. The National Conference on Criminal Justice (1973) has now articulated a set of standards which provide behavioral guidelines regarding humaneness and justice. Enforcement of these standards through the courts will undoubtedly have an impact on program operation and evaluation, for not only is the "right to treatment" at issue, but so are the general social and physical conditions under which the juvenile offender must exist. It is possible for a program to be effective (at least hypothetically) and yet not meet criteria of humaneness, fairness, and justice. The evidence, however, appears to be overwhelming that these three conditions are necessary to, but not sufficient for, effectiveness.

Behavior modification technologies and the conditions governing their use pose another issue to be addressed by those interested in evaluation of such programs. More and more juvenile correctional programs are instituting behavioral modification technologies and recently, electronic devices for remote observation and control of behavio have been proposed for juveniles, as well as adults, by Ingraham and Smith (1972), and by Schwitzgebel (1967). There is considerable debate about the use of these latter technologies, but the position paper by the American Friends Service Committee (1971) and the legal critiques by Shapiro (1972) and his colleagues in Southern California (1972) raise broad philosophical questions about freedom and psychic autonomy. Ingraha and Smith (1972) argue that electronic and other behavioral control procedures are preferable because the offender who accepts these devices has far greater freedom in other sectors of his life than if he were incarcerated. In rebuttal, Shapiro (1972) asserts that there is considerable risk of abuse of these technologies and that if behavioral control through electronics is instituted, we will have effected a fundamental change in values about personal privacy and freedom. Thus far, evaluators of behavior modification technologies in juvenile corrections have avoided dealing directly with these issues in their assessment of programs, but they will need to be addressed in the near future. Obviously the use of any interpersonal change technology raises ethical dilemmas regarding freedom of choice. The external manipulation of behavior should not be pursued without reference to the social values at issue. More than a decade ago,

Krasner (19 2) offered the following caution about the use of these technologies:

> Behavior control represents a relatively new, important and very useful development in psychological research. It also may be horribly misused unless the psychologist is constantly alert to what is taking place in society and unless he is active in investigating and controlling the social uses of behavioral control.

Our delineation of issues and dilemmas in evaluation in juvenile corrections would be incomplete if we did not refer to involvement of the offender population directly in the assessment of the program. Much evaluation is completed in corrections without any direct participation of offenders, and where they do participate, it is only to comment positively or negatively about the program as it is presented to them. They are not permitted to formulate expectations or suggest means for achieving objectives. This continues despite the suggestive data in the literature that one's expectations and assessments are critically linked to one's behavior. Certainly, this is the case when negative consequences occur. The youth who views his experience negatively, who is pessimistic about the future and his behavior is not likely to succeed in post-program situations. Youth are committed to correctional programs for vastly different behaviors and attitudes -- yet much evaluation starts from the assumption that there will be similar responses among offenders given a standardized program experience.

PROGRAM GOALS -- CAN THEY BE MEASURED?

Mott (1972) and others have proposed that effectiveness is to be defined with reference to the extent to which organizational goals are attained. Difficulties arise because of the ambiguity of the goals of corrections agencies, because some goals may be covert, and because there may be actual or potential contradictions among multiple goals. For example, protection of society is a typical goal in a correctional agency, but if it is carried to an extreme state, it would mean that an agency would seldom release an offender who was a "risk". Not to do so, however, would jeopardize the goal of rehabilitation.

The researcher, in his effort to measure organizational goals, must distinguish between ultimate and process or intermediate goals. These objectives must then be

linked to each other as we suggested earlier and they must also be measured in relation to the behavioral outcomes for juvenile offenders. Implicated in any effort at delineation and measurement of organizational goals is the definition of the problem that is or will be the target for change. The way in which the problem is defined influences the means of intervention for resolution of the problem. Thus, if delinquency is defined primarily in terms of the individual person rather than in relation to the peer group, neighborhood, or other social situation, the goals, targets, and technologies should relate to individual level phenomena. On the other hand, if the problem is viewed as situational, the physical social or economic environments would be the targets for change.

The use of goals in evaluation is also complicated by lack of clarity in both definition and in measurement techniques. Objective procedures for measurement of organizational goal attainment are presently not available for human service organizations. The research, therefore, must examine official mandates, objectives and priorities of the executive cadre, and staff perceptions of what should be the desired ends for the organization. The latter approach was utilized by Ullman (1967) in his study of the relative effectiveness of mental hospitals. He developed two normative measures of outcome which were independent of organizational definitions of the phenomena. The normative approach was the basis for recommendations of the President's Commission on Crime, and in the recently published Criminal Justice Standards of the National Advisory Commission on Correctional Goals and Standards (1973). The general assumption underlying these normative approaches is that the primary function of juvenile corrections is the rapid reintegration of the offender into his usual social roles with as little severing as is possible of ties with the external environment throughout the period of intervention and/or incarceration. Given this assumption, goals of custody and protection are of secondary importance and should so be assessed in evaluation. In the case of juvenile corrections, there is widespread consensus that the primary mandate is rehabilitation, but when staff perceptions are measured custodial ends often are asserted to be of equal importance. Thus evaluative procedures have to include multiple measures and variable weighting of goal priorities if these are to be utilized appropriately. The analysis of

six programs in the following section highlights some of the many problems involved in defining and measuring goals and in tracing out the consequences of goals for program performance.

A Critique of Six Studies in Juvenile Corrections

Six studies which deal with evaluation in juvenile corrections were selected to illustrate a variety of settings, programs, and populations, and to highlight variable questions and modus operandi for this kind of scientific endeavor. These selected studies are not representative of the entire range of research in this area, nor do they necessarily represent the most fruitful work. Rather, they are examples whose strengths and weaknesses can serve as guidelines for further refinement and elaboration of evaluative methodologies.

These six studies encompass a selection of research in juvenile corrections over a fifteen-year period -- research which looks at a variety of ways of trying to effect change in youth in a variety of settings in various regions of the United States. The researchers were involved in different ways with both the subjects and the settings in which their research took place. To assess this research, a framework was developed which emphasizes process components of effectiveness studies more than outcome measures, because it is the former which have received less attention at the organizational level and it is these which must be manipulated in programming.

First, the subject population -- what personal and social background characteristics were measured, what attitudinal, behavioral and personality charcteristics were assessed, and how were subjects selected. If there was both an experimental and control group, how were they compared and were they adequately matched? To what extent were these differences in the subject population related to later differences in outcome?

Second, the setting in which the intervention took place -- the milieu and situational characteristics, the characteristics of staff both in terms of background and attitudes, the goals of the program, the inter-organizational network in which it functions and the constraints upon it in general in terms of realizing program

goals. To what extent are these programs differentiated along organizational dimensions, and if they are, how are these differences considered in the overall evaluation of the program?

Third, the treatment technology being used on the subject population. Are the intervention strategies or technologies being used clearly and precisely specified and are there real differences for the control and experimental populations? Can we separate the effects of various segments of the intervention process? Is the technology consistent over different periods in the history of the project and, if not, are these differences taken into account?

The next step is to examine to what extent they sought to discover the effort expended to effect change. One cannot really compare technologies or programs with regard to their intrinsic value if they are not implemented with the same degree of intensity and if they are differentially embraced or resisted by their subject population. To what extent are the efforts of staff, the efforts of subjects, the efforts of the organization, and the efforts of the community analyzed in relation to a discussion of the technology and its effects? If there are differences between the experimental and control programs, are differences in effort variables looked at as possible explanatory factors?

Process variables are defined as the effects produced during the period of intervention. These include the effects on a short-run basis, i.e., prior to termination or at the point of termination from the program. For reasons asserted earlier with regard to recidivism, process effectiveness measures are assumed to be of crucial importance in evaluation and should not be viewed as inherently "softer" data than long-range effectiveness measures. Along with traditional types of process measures such as offenders' attitudes toward self, expectations for success on the outside, and views toward peers and staff; other behaviors engaged in during treatment are also considered including: objective measures of achievement while in the program and at point of termination, and preparation for future roles given by the program. The extent to which researchers examined humaneness, fairness, and justice for the subject population will also be considered. Also examined are length of

stay, costs of program, and proportions of subjects completing the program, failing, or dropping out, as well as process effects on the staff and organization.

Finally, the outcome or performance measures are analyzed as is the extent to which these are related back to the other aspects of the analysis. For example, to what extent are differences in recidivism related to differences in subject characteristics and the effort expended by the subjects during treatment, as well as to differences in the treatment technology? In what different ways is recidivism measured and how adequately? To what extent are other measures utilized, such as stability and survival of the program itself, morale and quality of the staff, degree of community cooperation and concern with the program, as well as behavioral measures of the subject's long-range adjustment through vocation and education.

"Social Structure, Identification, and Change in a Treatment-Oriented Institution," Raymond J. Adamek and Edward Z. Dager. 1968.

Characteristics of subjects. The subject population consisted of the universe of girls resident in the institution and data are given about the variable percentages of these girls with respect to several background measures. No comparative information provided about the distribution of these background characteristics for other institutionalized female delinquents. The institution had a selective admission policy with respect to age, educational level, and severity of emotional disturbance.

Subjects were also characterized by their entering scores on the ICL, the MMPI, and the IPAT Anxiety Scale, which were administered by the institutional staff at intake. These data were used without reference to the situational characteristics which might have affected responses.

Only five subject characteristics -- intelligence quotient, social class, religion, age at entrance, and length of stay were analyzed for relationships with degree of change. They found that length of stay and age at entrance were significantly related to change and degree of identification with staff, so these variables were controlled in the analysis.

Intervention setting. The intervention setting was described in terms of its location, auspices, authority structure, and staffing, and structure and mechanisms of social control as they affect milieu. These data were apparently gathered through observation, analysis of institutional documents, and questionnaires to residents. Goals of the institution were not clearly identified nor was there any indication as to how they were operationalized; thus effectiveness cannot be measured With respect to the environmental context, only sketchy information was provided and there were no data about the opportunity for or frequency of contact with local community people. Information is needed about the extent and degree of available outside contact and the degree of congruence between the intervention setting and the usual social roles to which residents will return, for these are important in accounting for the success or failure of efforts to integrate offenders into conventional societal roles.

Treatment/technology. The most distinctive aspect of the treatment technology was that it was highly-structured, unambiguous, and consistent in the patterns of reward and punishment. It was described in terms of its mechanisms of social control, but no information was given about sanctions or rules. Also not presented was information about the relative emphasis of the individual or group in treatment, the problem focus of staff, the kinds of diagnostic categories or classification schemes used, and the variety of techniques employed.

Effort. Information about contact with group mothers was provided, but there was no data on frequency of contact with social workers, psychologists or psychiatrists, number of hours spent in school, emphasis placed on group or individual punishments and rewards, or even the kinds of contact and the quality of contact of staff with girls. Two effort measures of students -- service as a group leader and conformity to institutional norms -- were related to the degree of indentification with the institution and staff -- a process variable. It would have been helpful had the authors also analyzed the relationship between those effort measures and outcome variables but this was not done.

Process. Process measures of interest to the authors were: identification of students with the staff and institution, and degree of change in self-esteem, faith in people, psychological-behavioral adjustment, and anxiety. They were not interested in any behavioral measures.

In accounting for the development of identification with the institution and the staff, the researchers alluded to a number of other effects of treatment, or process effects, which were not measured but were described such as: reduction of peer interaction, little or no peer support for "fighting the system," isolation from the outside, too much punishment for minor rule infractions, no legitimation of many of the rules by the girls, and so forth. Several of these potentially negative process effects required systematic study.

The measures of change were calculated for girls at various stages in the treatment process, i.e., they were not measured at point of termination, but rather all were obtained at the same time regardless of the phase of treatment. As indicated earlier there was almost total reliance on psychological assessment as a measure of change. Although the average length of stay was 18 to 24 months, information was not given about what types of data were considered in release decisions. No information was provided about changes in school performance, in peer relationships, in misbehavior, and so forth. Also unknown is the proportion of girls in this program who completed it, dropped out or were transferred. No information was given about program costs.

Outcome or Performance. This study did not report any outcome or performance measures either at point of termination or after termination. All change considered was change within the institution during the process of treatment. But, there was no certainty that any of these indicators of process change were linked to any long-term changes for the subjects, either positive or negative.

A Follow-up Study of Boys Participating in the Positive Peer Culture Program at the Minnesota State Training School for Boys: An Analysis of 242 Boys Released During 1969, Minnesota Department of Corrections, June, 1972.

Characteristics of Subjects. The subject population consisted of the boys

released in 1969 from the Minnesota State Training School for Boys. All the boys were characterized on numerous dimensions of background, but for only three of these characteristics -- area of residence at admission (urban/rural), racial/ethnic background, and intelligence estimate were any comparisons made between this Training School population and the male population in that state between the ages of 10 and 19. And yet, comparison after comparison was made in these figures between boys who had their paroles revoked from those who did not. It seems likely that they collected information on every background characteristic that appeared in the file without any theoretical reason or conceptual framework for linking them to the dependent variable or measure of effect. This resulted in a number of comparisons between the background characteristics of boys whose paroles were revoked with those whose paroles were not revoked being made without any really sound reason. Therfore, it was not surprising that some significant relationships were observed. A number of the background characteristics which were analyzed were never operationalized clearly, e.g. "living situation" and "drug/alcohol and physical abuse".

Several intra-institutional program characteristics were also studied in relation to the subjects -- that is, characteristics of the subjects that were related to their institutional experience. These included: the number of successful and unsuccessful truancies of boys from the institution, the differences in lengths of stay, the age at release from the institution, the school grade placement at release, the cottage lived in, and the living situtation after release as related again to revocation or non-revocation of parole. As mentioned earlier, the reader was not informed as to how the data were collected, nor how variables were defined and operationalized.

This study relied almost completely on the characteristics of subjects for the explanation of effectiveness or ineffectiveness of the program technology. These characteristics of the subjects were related not only to the incidence of parole revocation or non-revocation, but also to comparisons between parolees and transfers from the institution.

Intervention setting. There was virtually no description or analysis of the

intervention setting. It would have been helpful if the cottages which were analyzed in terms of the differences of the characteristics of the residents, had been describe in terms of any organizational or structural characteristics. It is critical to an adequate understanding of the process of treatment to determine the significant characteristics of the settings in which it occurs.

Treatment/technology. According to the report, the technology employed in treating these boys was Positive Peer Culture (PPC) -- a type of group therapy, based on principles of guided group interaction, which utilizes peer pressure and staff guidance to treat delinquent youth. The description of PPC states that it was the basic treatment tool of the institution. Group meetings five nights a week were the focal point of the boys' daily activities. The group was also the focus for many other activities experienced by the youth in the institution.

This is fine insofar as it goes, but it leaves a great many questions unanswered. What kinds of behaviors or attitudes did this technology really try to alter and what kinds of behaviors and attitudes did it accept or tolerate? To what extent were group meetings supplemented by other kinds of counseling such as individual counselin To what extent did the group reward and/or punish its members at times other than in group meetings? To what extent did all staff really implement the PPC program -- was it thwarted at all by custodial staff or other treatment programs? What particular parts of the PPC technology were crucial for effective treatment and what parts of it were adjuncts whose modification would leave the program relatively unchanged? Were all boys in the institution really subject to the same technology or were there differences in its application, intensity, or staffing patterns that might, in fact, have been related to the success or lack of success on parole? There was a difference in the parole success rate of different cottages in the institution -- it is probable that some of this might have been explained by differences in the application of the technology in these different units. No effort was made to look at the actual implementation of the technology in the institution as a whole, in different sub-units of it, for different kinds of offenders, and by different kinds of staff. This is a crucial deficiency in a study which seeks to

assess the effect of one kind of technology in a single institution. At the very least, it is necessary to examine the operation of that technology in depth in order to ascertain why it was or was not effective.

Effort. Ideally, the effort expended by staff and boys in implementing PPC should be considerable -- after all, "they have the responsibility to help and care for each other 24 hours a day". They met as groups five nights a week for 90 minutes, but no information was provided on the total hours of contact with staff or with each other in discussing and working on individual or group problems, how much in-service training was given to institutional personnel in PPC, or how much of a commitment the institution itself gave to the program.

No data was provided about the relative priority of the PPC technology in the regular operation of the institution, nor about supports provided by the institution for re-entry under a PPC technology, and also none about the effort and commitment by boys to the PPC program. Such differences in such commitment and effort might well be related to differences in the success of parole outcome and probably were related to the selection of boys for transfer as opposed to those who completed the program.

Process. The process measures included the number of truancies from the institution, the success or failure of these truancies, academic school grade placement at release, and the placement or living situation at release. These measures were not only insufficiently defined but were also not easily comparable with other kinds of programs. Further, they were always analyzed for their relationship to the outcome characteristics of the subjects -- whether or not they had successful parole outcomes -- but they were not analyzed for possible relationship with personal and social characteristics of the subjects or to differences in treatment.

Although information could have been obtained about attitudes and behaviors assumed to be influenced by the PPC technology, none of these were, in fact, obtained. For example, they could have had pre and post-program evaluation of school attendance and performance, employment expectations and behavior, attitudes toward delinquency, peer and self, or relationships with parents and relevant others. The only

intra-organizational variable which was evaluated was transfer status. Transfers were compared with non-transfers with several important differences noted. Twenty-three boys were transferred to other juvenile correctional facilities because of repeated truancy and behavior which required closer supervision and control. But, these transfers had fewer incidents of all types of truancy prior to admission than did the parolees. Why then did these boys run more than other boys at Red Wing? What possible interactions might there have been between the program and the characteristics of the boys that produced a propensity to run? These questions remained unanswered.

Transfers were, on the average, younger and had a higher estimated intelligence level than parolees. A higher percentage of the transfers were from the Metropolitan area and were minority group members. Moreover, there were fewer known cases of alcohol, drug, and physical abuse among the transfers than among the parolees. Why these particular boys were not able to complete the program was not answered despite the differential effect of this program on boys. Rather than simply stating, as the researchers did, that "certain strengths and characteristics may be necessary for a boy's gainful participation in confrontive peer group treatment programs,""it would certainly seem necessary to try and relate process measures to characteristics of the organization and treatment program, as well as to individual characteristics of participants.

Outcome or performance. The outcome measures used in this study were all variations on the general theme of recidivism. Of the 219 parolees, 51% had their parole revoked while 49% did not, and the comparisons between these two groups and the transfers were made in terms of the previously-mentioned background characteristics and intra-institutional characteristics. Revocations of parole were more frequent for members of racial minority groups, especially American Indians and Spanish Americans; for boys with more disruptive living situations; for younger boys; and for boys with more frequent truancies. In addition there were more revocations of parole of boys released from certain cottages and for boys with shortest lengths of stay in the institution. Along with this analysis, they collected data on the average numbe

of months on parole before violation (6.8 months), the offenses resulting in revocation of parole, and the institution to which the revocations were returned. No other measures of outcome were attempted.

Girls at Vocational High: An Experiment in Social Work Intervention, Henry J. Meyer, Edgar F. Borgotta, and Wyatt C. Jones. 1965.

Characteristics of Subjects. The subject population consisted of 400 girls who entered Vocational High between 1955 and 1958, and who had been identified as potential problem cases by the research staff in their examination of the school records. From this pool of four entering groups of students in four different years, a random procedure was used to select the girls for the experimental and control groups. Comparisons were made not only between the experimental and control groups on a variety of background and social characteristics, but also between the whole group of identified potential problems and the remainder of the high school populatio of the school.

Differences were observed among these girls identified as potential problem students and others on a number of dimensions, most of which were in the expected direction. The data indicate that the random procedure for selecting experimental and control cases among the potential problem population resulted in generally simila groups, at least on those variables for which comparisons were made.

Despite elaborate procedures devised for tapping background and social characteristics, these characteristics were, for the most part, not used in the later analysis of effectiveness and content of service. The reason for this we do not know, because many of these characteristics might well be related to outcome and would present a more detailed and informative picture of the actual process of intervention and its effects. For example, individual therapy might be more effective for certain types of girls and group therapy more effective for other types, perhaps depending on racial background, sibling structure, or personality type.

Intervention setting. All subjects attended a vocational high school in New York City which had an enrollment of about 1,800 students admitted from all over the city. Subjects who had been randomly assigned to the experimental group were

referred to Youth Consultation Service, a private non-sectarian social agency for individual casework and group therapy.

During the evaluation, a decision was made to switch from individual to group treatment because of evidence that it would prove to be more satisfactory. The approaches differed in their settings -- casework in the traditional agency setting and groupwork in a more relaxed community setting . Although this was not ever explicitly stated, the differences in the setting alone could have been important in accounting for the differential impact of the two treatment modalities.

The goals specified for intervention were vague and not well-defined. They essentially revolved around trying to interrupt potential deviant careers and assumed that the agency was successful in diagnosing potential problems. Treatment objectives appeared to be multiple and individualized, so there were very real problems in defining effectiveness criteria. There was no information as to how many regular staff members were involved in this experimental program, the actual ratio of staff to girls in various phases of the project, the background characteristics of the staff that might be related to their differential effectiveness in working with these girls, or the relative priority of this project in the on-going functioning of the agency.

Treatment/technology. Originally, the project sought to use casework services on an individualized treatment basis as the primary technology, but in the second year, they shifted to a process of group referral and treatment. These two technologies were compared and evaluated along with the comparison between the control and experimental groups. Unfortunately, there was a wholly inadequate presentation of the differences in the two technologies and how they were actually implemented. There was a much better description of the ways in which the group treatment processes operated than of the ways in which the individual treatment processes worked, probably because group treatment was seen as more novel and interesting in this setting.

Effort. Fairly complete information was provided about the efforts made by the school, the agency, and the girls to implement and maintain this program. Effort variables on the part of the school and agency included: provision of physical facilities; scheduling changes; orientation of staff and shifting of resources, staffing

patterns, and treatment strategies. For staff, there were measures of contact with parents, number of interviews and content and depth of group sessions. In addition, the data indicated that, for the whole experimental group, 95% received some treatment services and half of these had 17 or more treatment contacts with social workers. Therefore, the experiemental group was clearly well-exposed to the therapeutic program In addition, girls in the experimental group reported more help from social workers and researchers than did those in the control group. The researchers were sensitive to their impact on the situation. Other effort measures collected for the experimental girls included judgements by the caseworkers and group therapists about their clients. The caseworkers felt that not many of their clients became seriously involved in a treatment relationship on an individual basis, but this was not the case for girls involved in group treatment. Unfortunately, the girls themselves were not asked to rate their own effort, and we have no information on the number of missed appointments or instances of late appointments.

Process. Because of the lack of precision in defining goals and technologies of this program, the researchers were confronted with a real problem in assessing effectiveness. They tried to resolve this dilemma by presenting an array of variables designed to measure the impact of service, all of which were variants of process measure These process measures included:

1. Judgment by the caseworker as to the progress made by the girl in using the agency's services constructively. Background variables of the girls (race, religion, intelligence, and clinical diagnosis) were examined to determine whether they related to the degree to which the girls were involved in using such service. Unfortunately, there was no examination of the background characteristics of staff, as related to their ability to motivate or help their clients.

2. Judgment made by the caseworker as to the effects of treatment. The Hunt-Kogan Movement Scale showed more positive results for group treatment. For the experimental sample, only one-fifth of all the girls were judged to have changed or moved positively during treatment.

3. School status at the end of the project. Twenty-nine percent of both the experimental and control cases had graduated high school at the termination of the project. Equal proportions had dropped out. There was no discernible impact of the intervention on this measure.

4. Highest school grade completed. Extremely small differences in staying in school favor the experimental cases.

5. Academic performance. There was a positive selective effect of the treatment program in reducing failing grades in academic subjects.

6. School-related behavior. There were no significant differences in attendance records. There was slightly less truancy among experimental than among control subjects. There was no difference in conduct marks between the experimental and control subjects. There were no real differences in teacher ratings on character and work traits. There were no real differences in ratings by guidance and counseling staff between the experimental and control subjects.

7. Out of school behavior. There were no real differences in entries on health records, or in instances of out-of-wedlock pregnancy.

8. Client self-reports of effects. There are very scant differences, if any at all, on measures of well-being, perception of quality of interaction with others, psychological insight and reactions to help, or assessment by seniors of their present and future situations.

9. Personality tests. There were no real significant differences in responses on the Junior Personality Quiz or the Make a Sentence Test.

10. Sociometric measures. There is no evidence of effect.

Although there is an impressive array of these process measures which are analyzed, many were not systematically collected and were not really meaningful for purposes of analysis. Rather, they were interesting observations about adolescent girls and agency staff. Unfortunately, there were no process measures designed to look at the effect of this experimental program on the agency staff or the school. Such would seem to be almost as important as discerning the impact on the client if the program were to be generalized to other settings.

Outcome or performance. There were no long-range outcome or performance measures used in this study. It was apparently sufficient to show that even process or intermediate objectives were not fulfilled by the intervention strategy. But a study which sought to evaluate a delinquency prevention program would seem bound to at least try to assess a few long-term effects.

The Silverlake Experiment, LaMar T. Empey and Steven G. Lubeck. 1971.

Characteristics of subjects. The subject population consisted of 261 boys who were assigned by the courts from a common population of delinquent offenders in Los Angeles County.

Once boys were selected for the project, they were randomly assigned to either the experimental or control program and the background characteristics of the subjec

in these two programs were compared on a host of background variables. Few differences were observed, but the experimental program may have inadvertently been assigned more serious or more experienced offenders.

Many of these measures of background characteristics were combined into scales and were subsequently used in analysis of predictor variables for both process and outcome differences. Thus, the scales of peer influence, background, offense, and personality were consistently used throughout the report in explaining instances of critical incidents, runaways, program failures, and program successes.

Intervention setting. There were very clear distinctions in the setting for intervention for the control and experimental population and these differences were considered among the most crucial elements of the intervention strategy. The control program was housed in a large physical plant of the traditional institutional variety while the experimental program was located in a ranch style house in a residential neighborhood. The two programs also differed in size; no more than 20 boys at a time in the experimental program but up to 125 boys in the control program. Boys in the experimental program returned to their own homes on weekends and attended high schools in the community while boys in the control program lived there all the time and attended high school on the grounds.

Actual differences in the staff-inmate ratio in the two programs were not given, but the quality and quantity of the interaction between staff and offenders differed between the two programs. A special questionnaire was constructed and administered to both the staff and boys in both the experimental and control programs. The instrument dealt with the perceptions of the setting, among other things, and it verified that there were indeed differences in the expected direction between the experimental and control programs.

The two programs differed considerably, not only in the flow and content of communication between staff and offenders, but also in the degree of community linkage with other agencies concerned with youth. The rule structure and sanctioning mechanisms also varied in accord with traditional differences between custodial and treatment-oriented programs.

Treatment/technology. The authors reported that the control was highly structur and traditional with much emphasis placed on training and individual counseling. Custodial factors were of moderate importance. Far more detailed information is provided about the experimental program, about theories of causation, treatment modalities, and so forth. The basic technology was a variant of "Guided Group Interaction" The peer group was utilized as both the target and the medium of change. The second important component of the technology was community linkage, providing a way in which offenders could be reintegrated into normal community roles. The experimental unit was located within the community; youth attended neighborhood schools, and had weekend home visits.

During this research, both programs underwent a series of major organizational changes which were reflected in changes in treatment methods. The experimental program imposed a strong negative sanction against runaways sixteen months after it began. The control program shifted from a concentration on one-to-one treatment to more group oriented modes and to more participatory forms of decision-making by both staff and boys. Both sets of changes were analyzed in terms of their relation to changes in process and outcome efforts.

Effort. The amount of effort put forth by the staff in implementing the program cannot be determined accurately, because sufficient information was not given, but one got the impression of fairly high levels of effort in both the experimental and control programs. More attention was paid to the efforts of boys to embrace or resist the different programs by asking them through interviews, questionnaires, and informal conversations about their own and others' involvement in the treatment process. Information was not provided, however, on the amount of effort put forth by boys in the school (there was no comparison of attendance records) at at work (there were no records of work attendance or performance).

Process. Some interesting behavioral measures of the effects of treatment on participants were utilized, including runaway rates, participation in critical incidents, and program failures. The terminal runaway rates for both programs, covering a two and one-half year period, were relatively high and very similar -- 37% at the

experimental program and 40% at the control program, but these rates were associated with different predictor variables in the two kinds of program.

Critical incidents were studied only in the experimental program, so cross-program comparisons of participation were impossible. This type of analysis is important in correctional program evaluations because it is an enlightening process measure, not only in delineating the types of concerns that are defined as crisis-provoking for the program, but also in terms of organizational handling of crisis and the participants' responses to it.

The program failure rates (those transferred or drop-outs) of both programs were much lower than the runaway rates. In both programs, failures were boys with the most persistent and serious offense backgrounds and personality problems.

The summary measure of the process effect of treatment in the two programs was that no more than 46% of the experimental subjects, as contrasted to 50% of the control subjects, successfully completed the program. The remainder were runaways or in-program terminees. Surprisingly, there was no information collected from the participants in either program at the point of termination or before with regard to their own evaluation of the experience or their assessment of the impact of it on their futures.

This study dealt rather thoroughly with some of the negative effects of treatment on participants. Two particular aspects stand out: (1) There was evidence that the experimental program may have been overly concerned with custody and control (2) both programs were far more inclined to punish boys for undesirable behavior than to reward them for desirable behavior.

This study was one of the few which compared the costs of the two programs and the average lengths of stay. The average monthly cost for the experimental program was $302.86 per boy, while it was $362.18 per boy for the control program. But the average length of stay for experimentals was 5.73 months as compared to 12.65 months for controls; thus, the difference in cost was considerable. The human costs for the boys were also less, for they had much more contact with home and their usual environment with shorter periods of confinement in the experimental program. Costs

were not examined with respect to relative differences in humaneness, fairness, or justice between the two programs, but this might have been part of this type of process evaluation.

Outcome. Outcome was analyzed only in terms of recidivism in this study, but the ways in which it was measured far surpassed the usual recidivism reports. All experimental and control subjects -- runaways and failures, as well as successful graduates -- were followed for at least a year after their termination from either program. Recidivism was analyzed in three different ways: (1) in terms of number of offenses committed by individuals; (2) in collective terms by comparing the total volume of delinquency committed by subjects a year before entering the program with their total volume committed a year after termination from the program; (3 in terms of the seriousness of the recidivistic offenses. This relative estimate of recidivism certainly has advantages over the more traditional absolute approach.

Most subjects in both programs seemed to have been relatively free of post-program delinquency and there was relative similarity in the amount of individual recidivism from both experimental and control programs. Subjects who did not complete either program were much more likely to recidivate than those who did. There was a 73% reduction in the volume of delinquency committed by the experimental subjects and 71% for control subjects during the twelve-month period after release, as compared to the twelve-month period before assignment to the program, but these findings are possibly due to regression effects. Both programs brought about a significant reduction in the volume of serious offenses, suggesting that they were important sources of delinquency control. Although there was no assessment of long-term outcomes, these data do suggest that the experimental program, which was much shorter and thereby less costly, was at least as effective as the control program.

The Youth Center Research Project, Carl F. Jesness, William J. DeRisi, Paul M. McCormick, and Robert F. Wedge. 1972.

Characteristics of Subjects. The subject population consisted of 904 wards of the California Youth Authority, who were randomly assigned to the two programs -- O.H. Close School which used transactional analysis or Karl Holton School which used

behavior modification. These boys were compared on several dimensions of individual and background differences, but few differences were observed between the study populations at the two schools. Despite the impressive array of data that were collected, none of these measures were later analyzed for their relationship to measures of process or outcome effectiveness, with the exception of the personality classification measures.

Intervention setting. A primary objective of this research was to examine the differential impact of two technologies in two institutions, apparently alike in their organizational structure, staffing patterns, and physical layout. Throughout the period of the study, the organizational structure and the number and the types of personnel were almost identical. The existing treatment programs at the two schools were also almost identical, at baseline. The internal climate of the two schools was described in detail, but little information was provided about the environmental setting in which they existed. They were situated adjacent to each other in Stockton, California, but there was no specification of how isolated they were from the rest of the city, how much communication existed between the two schools, how much interaction there was between staff and boys from the two schools or the network of community resources that was linked to the schools.

Treatment/technology. The heart of this research was its differentiation and comparison of two treatment technologies -- Transactional Analysis and Behavior Modification. The two technologies were presented as somewhat idealized conceptions of what should actually occur in their implementation, but differences were apparent in the goals for the two technologies, the content of in-service training sessions for the two sets of staff, the composition of treatment groups, the ways in which participants were rewarded and punished, the expected actions and routines of staff and boys under the two systems, and the expected changes in behavior and/or attitudes. The two strategies were supposed to be particularly effective for a broad range of behavior and personality.

Although the written conception of the operation of the two technologies varie there was no indication that these differences pervaded all areas of institutional

life. The boys in both programs went to school with their own hall group, and were taught in an individualized, non-competitive manner. Both technologies involved heavy use of group methods, contracts between staff and boys, emphasis on improved social adjustment, and a decrease in the probability of delinquency. Readers would be aided by having descriptions of instances of differences in handling of the same behaviors, or differences in the kinds of rewards given for the same behaviors. More clear-cut differentiations in the actual operation of the program components were necessary to really understand that the differences in the two technologies were more than a matter of differences in jargon.

Effort. The extent to which either or both of the treatment technologies were actually effectively implemented is debatable. There were apparently quite serious problems in conveying to line staff that these technologies were more important than the issues of security, paperwork, or housekeeping. The priority for the treatment program within each institution might have been different for there was data to suggest that Transactional Analysis was actually implemented more successfully than Behavior Modification. There was no continuing in-service training for the staff in Behavior Modification, but there was continuing training for Transactional Analysis.

The researchers also took considerable pains in measuring the effort of the subjects to embrace or resist the treatment technologies. These measures of effort of participants were then related to their maturity and ego levels. For all these measures of effort, it was determined that staff and boys of both schools approximated the original expectations of the project, but fell short of the ideal.

Process. One of the principal process effects examined was change in management problems within the institutions. Both programs, had an eventual drop in the number of incidents of misconduct reported, although initially the number rose. They also discovered that there was a continuing reduction in the number of residents sent to detention for misconduct, and when it was used, it was used for briefer periods and "for promoting change in behavior rather than for retribution." However, since these data were based on special incident reports written by staff in both institutions they

may not be comparable. This indicator was more useful for examining trends within a single unit than in comparing two units.

Also examined were several psychological indices of process effects. Results from the Jesness Inventory and the Post Opinion Poll showed positive psychological change in both units, but again Transactional Analysis subjects evaluated the program more positively. Observer ratings of behavior indicated greater behavioral change in the other unit as might be anticipated.

There were many hypotheses which sought to relate the characteristics of subjects to the process effects of the different kinds of treatment and they were as follows:

1. Transactional Analysis will be more effective with higher maturity subjects, and this will be evident in their verbal behavior and in their observable behavior.
2. Transactional Analysis will be most effective with higher maturity subjects who enter treatment expecting to change, and who receive high intensity and high quality treatment.
3. Behavior Modification will be more effective in changing the behavior of lower maturity subjects.

None of these hypotheses were clearly substantiated. Indeed, they found that the more mature subjects did better in both treatment programs.

Unfortunately, negative process effects were not examined for either program. There was a consistently longer average stay for subjects in the Behavior Modification program than for those in Transactional Analysis. No information was provided about cost differences of the two programs, although one would assume that Behavior Modification was more expensive since it had a longer average length of stay.

Outcome or performance. Parole revocation was the only long-run outcome measure used, as was true in many of the earlier studies. They compared the subjects' rates of parole revocation for a 12-month period following release with those of inmates from the same institutions who had been released prior to the introduction of the treatment technologies and with inmates from two other institutions. It is important to note here that they only looked at the rates for successful graduates, not for

runaways or in-program failures, and they did not compare these rates with the prior delinquency of the subjects as was done in the Silverlake Experiment.

In a 12-month parole exposure period, only 31% of the Transactional Analysis subjects and 32% of the Behavior Modification subjects of the same age had been removed from parole, and these figures were significantly lower than those of the control groups. No other outcome measures were used so recidivism again reigned as the supreme criteria of effectiveness.

A Comparative Assessment of Probation Practices and Perspectives, Stephen Burghardt, Rosemary C. Sarri and Carl Gohlke. 1971.

Characteristics of Subjects. The subject population consisted of offenders who were assigned to probation in a large mid-western city; offenders in a new institution utilizing a differential treatment model and offenders in a traditional training school program.

All of these populations were compared in terms of the background and social characteristics. The probation population selected was slightly older, more urbanized, had more years of education, and had a higher proportion of Blacks than the institutional population. The probation population was differentially classified on the basis of personal characteristics for service.

Few significant differences were observed in the commitment offense characteristics of the institutional and probation samples, but the institutional population had a larger proportion of offenders with more offenses. This same distinction between the number of offenses committed also holds for three types of probation groups -- there was a direct relationship between the number of past incidents and the type of caseload - (i.e. minimal, normal, intensive) to which an offender was assigned. The fewer the number of incidents, the greater the chance of being placed on a caseload with minimal service interaction. However, for the most part, the data in the background and offense characteristics of the several populations were striking because of the similarities they revealed between the probation and institutional populations, not the differences. These data refuted the often-stated

assertion that probationers are less serious offenders.

The background characteristics of offenders were not controlled in the analysis of process outcomes, though this was partially fulfilled because of the interaction between disposition (i.e. type of probation caseload or institution to which the offender was assigned) and his background characteristics. One of the striking but tentative conclusions they drew from the analysis of the relationship of race to dispositional outcome was that Blacks from the same social class and of equal educational achievement as the whites on "intensive" caseloads were not assigned to probation at all but probably were "tracked" into institutions. Data were insufficient, however, to test this hypothesis.

Intervention Setting. There were marked differences in intervention settings for the three programs. Probationers all lived in a large, mid-western city on their own while the institutional populations were housed in large complexes in rural areas. The programs were characterized by placement on a continuum of restrictive to non-restrictive milieu, and this facet of the intervention setting clearly was the crucial distinction between probation and the institutions.

Treatment/technology. the Probation Office utilized the "San Francisco Plan" for classification of offenders and caseload assignment. Using the criteria of age, probation offense, psychiatric test scores, and prior offense record, probationers were placed in minimal, normal, and intensive caseloads so that needs were matched with services. The actual strategies of intervention with probationers were not, however, systemmatically explicated. Service and the ways in which it was to be given, the content of discussions and the techniques of treatment were not reported sufficiently.

Both institutions functioned under similar general guidelines on training and custody, but there were significant differences in their technologies. The significant differences in the actual and desired goals for the institutions perceived by the staff were reflected in clear differences in the following dimensions of treatment technology: emphasis on process goals of custody and control, and treatment/community

involvement as opposed to product goals of education and rehabilitation; the degree of complexity and sophistication of the technology and the degree to which the technology emphasized interpersonal relationships between staff and clients or social distance between staff and clients.

Effort. Measures of the effort of staff included:

(a) relative amount of time spent in various aspects of service;

(b) offender and staff perceptions of the efforts of staff to care for and and help offenders;

(c) reports from both the staff and offenders of the frequency of discussion of personal problems;

(d) reports of probation officers and probationers with regard to effort at referral to schools, employment agencies, hospitals, vocational rehabilitation, and other agencies but no comparable information for institutional staff.

The only measures of offender effort included in this report were:

(a) Offenders reports about problem-sharing with probation officer, minister, family, friends, and so forth;

(b) Efforts of probationers to contact referrals.

Process. This study was primarily concerned with the effects of service on clients during and at the point of termination from the program. A large array of process measures were collected, including the following: attitudes of staff with regard to their interest and involvement in service technologies, their type of interaction with young offenders, and their attitudes toward offender change efforts; offenders' perceptions of their interaction with staff, their optimism about the future, feelings of self-esteem, and faith in people.

Probation staff had more positive attitudes toward offenders and their chances for improvement than institutional staff, and were more interested in factors related to changed goals for clients. The typical probationer-probation officer relationship compared favorably with the perception of interaction between staff and offenders in the institutions but significantly fewer probationers felt that they had been helped

in either job training or with matters relating to problems outside probation itself than was true in the institutional population.

The researchers also considered the question of the relative cost of the program. The estimates of the per client cost per day in the various programs is as follows: $15.90 for an inmate in the traditional unit; $30.00 per day in the new program; and, $1.09 for a probationer/parolee. Thus, there was little convincing evidence from the comparison of process measures for widespread institutionalization of young offenders. Data were insufficient for comparison about the relative lengths of stay in each of these programs.

Outcome or performance. The researchers did not utilize any long-term indicators. They relied exclusively on intermediate measures of effectiveness, referred to in the earlier sections, although allusions were made to the relationship of effective process outcomes and long-run impact.

A SUMMARY CRITIQUE

The adequacy with which the researchers in these six studies evaluated programs will be summarized in a series of small charts for each dimension followed by a brief discussion of the overall strengths and weaknesses of the analysis of these dimensions. The reader is cautioned, however, that these judgments are crude and somewhat subjective, and do not reflect the differences in objectives, units of analysis, problems, and resources available.

Subject Characteristics

Excellent	Good	Fair	Poor
Empey and Lubeck	Adamek and Dager --- Meyer, Borgatta, and Jones --- Jesness, DeRisi, McCormick, and Wedge --- Burghardt, Sarri, and Gohlke	Minnesota Dept. of Corrections	

For the most part, researchers handled the explication and measurement of subject variables adequately, though there was a tendency to measure a large number of them without any apparent theoretical reason for so doing. In most cases they were used primarily to show the comparability of the experimental and control populations but were not systemmatically analyzed for their relationship to the effects of treatment. For example, every one of the studies looked at race in comparing the experimental and control groups but race was never further analyzed in terms of the extent to which there were differential rates of effectiveness in process and outcome based on race of the subjects. This would seem to be a rather important omission in the study of programs which handle disproportionate numbers of minority group members. Finallyl because there weren't many efforts to compare the population characteristics of the subjects with those of other correctional populations, these studies are subject to the criticism of the inability to generalize from them to the correctional population in general.

Intervention Setting			
Excellent	Good	Fair	Good
Empey & Lubeck		Adamek & Dager	Minnesota Dept. of Corrections
Jesness, DeRisi, McCormick & Wedge		Meyer, Borgatta, and Jones	
		Burghardt, Sarri, and Gohlke	

Most of these studies suffered from lack of attention to the effects of the intervention setting on evaluation process and outcome, particularly in the areas of goal description and analysis and the environmental context of the program. In only two of the studies (Jesness et al, 1972 and Meyer et al, 1965) was any mention made of any changes in aspects of the intervention setting during the course of the research but we suspect that such change is fairly common in most settings and must be dealt with in designing such evaluative research designs.

Treatment/Technology			
Excellent	Good	Fair	Poor
	Jesness, DeRisi, McCormick & Wedge	Adamek & Dager	Minnesota Dept. of Corrections
	Burghardt, Sarri & Gohlke	Empey & Lubeck	Meyer, Borgatta & Jones

The delineation and description of the methods of treatment was clearly a problemmatic area although all authors described some abstract conceptions of technology. In many of them, there wasn't even any comparable information given for the experimental and control groups with regard to the relative amounts of individual or group treatment; the premises underlying the choice of the technology with regard to problem causation, types of subjects most amenable to treatment, and priorities for problem focus; or clearly stated treatment objectives. Moreover, none of them adequately handled the actual implementation of the technology with regard to differences in methods of social control, rewards, restrictions, content of treatment sessions, locus of authority and decision-making and backgrounds of treatment personnel. In two of these studies, there were very clearly reported shifts in the technology during the course of the research (Meyer et al, 1965 and Empey et al, 1971) and in both of these the results were clearly related to differences in process effects. This is an important step and future research designs should be flexible enough to incorporate such changes into the total assessment of impact.

Effort			
Excellent	Good	Fair	Poor
Jesness, DeRisi, McCormick & Wedge	Meyer, Borgatta & Jones	Adamek & Dager	Minnesota Dept. of Corrections
	Empey & Lubeck		
	Burghardt, Sarri & Gohlke		

Most of these studies related few measures of the effort of staff to the goals and technologies of the program. Usually effort meant the frequency of contact with subjects and/or the number of interviews, group meetings, or counselling sessions held. With few exceptions, there was no effort to relate these measures of effort to the eventual process or outcome effects. The attempt to measure the effort of the subjects was even more truncated, though one would assume that such effort should affect the kinds of individual effects that occurred.

Process			
Excellent	Good	Fair	Poor
	Empey & Lubeck	Adamek & Dager	Minnesota Dept. of Corrections
	Burghardt, Sarri & Gohlke	Meyer, Borgatta & Jones	
		Jesness, DeRisi, McCormick & Wedge	

There are huge gaps in the conceptualization and measurement of process effects in most of these studies, although a few of the studies used a multitude of them. Some of the studies used a variety of behavioral measures of process effects to the exclusion of the subject's own attitudes and evaluation while others proceeded in the opposite way. Few of these studies used any measures of the effects of the program on the staff or the agency with regard to staff morale, turnover, agency innovation, agency interaction with each other, organizational and community networks, and so fort More than half of them did not include any statement of the costs of the program though most of them did evaluate the length of stay involved. Although the positive effects of these programs were always emphasized, little attention was given to the negative effects.

Outcome			
Excellent	Good	Fair	Poor
	Empey & Lubeck	Minnesota Dept. of Corrections	Adamek & Dager
			Meyer, Borgatta & Jones
		Jesness, DeRisi McCormick & Wedge	Burghardt, Sarri & Gohlke

In half of these studies, there were no measures of long-run impact of the program on the offenders or program at all and in the other half, the only measures used were related to recidivism or parole revocation. The measures of recidivism which were used, with one exception, were absolute and did not take into account the past history of offenses committed by the subjects. In an earlier section of the paper we discussed the importance of relative measures if recidivism is used as an outcome criterion.

Because recidivism was used so extensively, there was virtually no attention paid to other possible negative long-run effects of these programs such as stigmatization, decreased educational and occupational achievement, and feelings of injustice and anger. Furthermore, there were no efforts to look at any long-run positive effects of these programs - such as, educational achievement, vocational training, increased positive interaction with family and friends. In many of the programs, because there were no discernible positive features of process, outcome measures were not attempted.

A PLAN FOR COMPARATIVE ASSESSMENT

Our present research involves assessment of a limited number of varying types of juvenile correctional programs in a number of states in all regions of the country (Sarri and Vinter, 1972). This is policy-related research which seeks to identify the range and variety of present policies in juvenile correctional programs throughout the country, their relative effectiveness and ineffectiveness; what new alternatives might or are being developed; and how specific change can be brought about. Effort is being made to measure critical experiences and outcomes for offenders in

each of these programs. Organizations, however, rather than individuals are the primary object of analysis so rations, rates and probabilities will be calculated for groups of individuals. Furthermore, the assessment focuses more on the events within the temporal and social boundaries of the organization providing service, although some attention is given to cross-boundary relations, offenders careers within organization-sets, and post-program behavior.

In developing a plan for identifying operational criteria, we considered the three targets identified earlier: program outcomes for indivuals and organizations; organizational processes and effort, and inter-organizational processes and effort. Correctional units were classified according to a typology of the major functions performed by different agencies within juvenile justice systems: prevention and social control (e.g. youth service bureaus and community diversion units); identifying and nominating youth as offenders (e.g., police and school referral units); processing and referring offenders (e.g., court intake, diagnostic services); adjudicating offenders (e.g., juvenile courts); containing and controlling offenders (e.g., detention facilities, jails, custodial institutions, some probation and parole services); treating offenders (e.g., some probation services, community-based programs, some rehabilitative institutions); and, re-entry for offenders (e.g., some parole services, work release, job placement, some ex-offender organizations). This typology facilitates differentiation between units having the same general labels, but which may employ contrasting technologies or whose intended purposes are clearly different.

Our more elaborate classification of the functional categories of juvenile corrections was collapsed into four major groups for purposes of assessment: detention programs, processing, change and control, and exit management or re-integration programs. This classification recognizes that some corrections service units employ a combination of programs for part or all of their offender populations. It combines organizations that may pursue contrasting goals because their functions, programs, and results must ultimately be compared.

Questions had to be resolved about the relationship of fairness, humaneness, and justice to effectiveness before operational criteria could be formulated, as we

mentioned earlier. The assumption was made that humaneness, fairness, and justice were essential preconditions for any effective program. Thus, the decision was made to measure them independently of other measures of effectiveness. A set of standards of the National Advisory Commission of Standards and Goals (1973) is being adapted for this measurement. Illustrative of the criteria for humaneness and justice in handling offenders are the following:

1. Adequacy of sustenance conditions.
2. Nature and scope of interaction with peers and family members.
3. Access to and use of community resources.
4. Due process and other protections of individual rights.
5. Degree of restrictions imposed on offenders.
6. Extent of segregation of facilities.
7. Extent of discriminatory handling due to offenders' ascribed characteristic
8. Provisions for insuring the right to treatment.

The criteria for evaluation which are presented in the following section involve both program outcomes and organizational processes. They are drawn from our review of the evaluation literature on correctional programs and from studies of other human service organizations. The availability of several comparative studies of juvenile corrections was a great asset in selecting criteria. Among these studies are the Pappenfort and Kilpatrick (1970) census of children's institutions; the Street, Vinter, and Perrow (1966) study of institutions for male delinquents; the Lerman (1968) and Bailey (1966) reviews of correctional outcomes; and several extensive reports about evaluation design, methodology, process, and outcome, including the work of Rossi and Williams (1972), Weiss (1972), and Caro (1971). Criteria used for selection included the following: (1) linkage to major policy questions mandated for the research; (2) relative ease of measurement; (3) potential for achieving operational comparability across units, communities, and states; (4) theoretical relevance and potency; and (5) observed empirical potency.

Program Outcomes. These refer to the degree to which organizations achieve their intended results at both the individual offender (target population) and organ-

izational levels. Since both manifest and latent (positive as well as negative) results are being observed, the ration between them provides one summary measure of relative effectiveness. Illustrative variables about which data are being collected for each major program category are presented below. These criteria are not meant to be comprehensive; in general, the research strategy calls for a parsimonious range and choice of data collection measures. In the case of each type of program the criteria are being operationally defined. This process will be illustrated here only for one type of program, "detention," but the process is similar for the other types. Detention as a type of program is contrasted from the others because effectiveness is measured with respect to restrictiveness of the custody, the length of stay, the number of youth detained, and the proscense of screening procedures, more than with respect to the quality of the detention care per se. This is not to indicate that quality of service delivered is unimportant, but to specify other important criteria for a unit that is only to hold youth for prodessing prior to adjudication or disposition.

Detention Programs.

1. Type of detention and average length of stay.

   The less restrictive the detention and the shorter the length of stay, the more effective the unit.

2. Proportion of persons detained who have hearings within 24 hours.
   The higher the proportion of hearings held within 24 hours, the more effective the unit.

3. Percentage of juveniles detained for felonies.
   The higher the proportion of the detained who are charged with serious felonies, the more effective the unit.

4. Percentage of juveniles who are subsequently institutionalized.
   The higher the proportion of those detained wh are subsequently institutionalized, the more effective the unit.

5. Offenders' perception of living conditions, the more effective the unit.
   The more positive are offenders' perceptions, the more effective the unit.

6. Staff effort to curtail the domain of detention.
   The greater the staff effort to limit the use of physically restricting custody and to develop other alternatives for assuring court appearance, the more effective the program,

7. Independent observers' assessment of detention programming. The more positive the ratings of education, diagnosis, medical care, and so forth, the more effective the program.

Processing Programs. For units primarily concerned with screening, adjudication, and referral, the extent to which dispositions vary and correlate with differences in offenders' social characteristics (including offenses) constitutes a general measure of outcome. Organizational effort and process variables have particular importance for these organizations because of their crucial role in decision-making regarding the status of youth.

1. Extensiveness of offender screening and diagnosis.
2. Congruencies between diagnoses and disposition outcomes.
3. Percentage of juveniles referred to non-isolating and non-stigmatizing service programs.
4. Nature and amount of service delivered per offender.
5. Proportions and characteristics of juveniles whose case-processing shows adherence to due process procedures.
6. Offenders' generalized evaluations of the processing and quality of services received.

Change and Control Programs. Although these programs are differentiated according to community location, goals, degree of institutionalization, etc., key variables are isolated for the assessment of each type of unit. Three facets for measuring outcomes are considered: (a) offenders' generalized appraisals of the program and organization; (b) nature and extent of preparation for offenders' reintegration into the conventional world; (c) relation between service received and subsequent recidivist behavior of offenders.

(a) Generalized offender evaluation of the program, its goals, policies, and technologies; functions, patterns, and structure of friendships and other informal systems; offender growth in self-esteem and self-knowledge; ability to handle stress and frustration; independence, and individuality; and relative optimism regarding future.

(b) Degree and type of preparation of offenders for reintegration into the conventional world.

1. Levels of educational preparation and achievement.
2. Extent and content of vocational training and levels of achievement.

3. Extent and scope of occupational experiences provided by the programs, and levels of participation.

4. Nature of preparation for return to family, peers, and community situations.

5. Behavior of youth in relating to conventional social roles (work release, passes, etc.).

(c) Recidivist behavior of offenders in relation to service received.

1. Extent and type of post-unit offenses by self-report with controls for age and pre-program offense history.

2. Duration of avoidance of new offenses.

3. Extent of integration into non-criminal social roles.

Exit Management and Reintegration Programs. Concrete examples of such programs are parole, aftercare, pre-release centers, etc. We expect to find fewer units providing such services and proportionately far fewer than for adult corrections systems.

1. Percentage of offenders recommitted to correctional programs.

2. Offenders' perceptions of and judgments about re-entry services received.

3. Offenders' self-reports of legally proscribed behavior.

4. Percentage of offenders in vocational or academic training programs.

5. Percentage of offenders employed.

6. Extent and frequency of staff intervention in the community on behalf of offenders.

7. Average amounts of service received by types of offenders.

8. Extent of offender-staff planning of the reintegration process.

Organizational Provesses and Effort. Measures of phenomena or results at one level of assessment may be analyzed as measures of effort or process at another level. In studying organizational processes and efforts, we will attempt to identify efficiently those elements that contribute most to units' intended results or purposes. Again, it is assumed that organizational processes may have either positive or negative consequences, or both, for these ends.

Organizational Goals

1. Content and specificity of goals.
2. Priorities among multiple goals (e.g., treatment, rehabilitation, staff morale, custody, etc.).
3. Staff commitments to goal priorities.
4. Offenders' commitments to the organization's goals and their priorities.

Executive leadership.

1. Executives' goal priorities and commitments.
2. Nature of executives' relations with external units within the juvenile justice system, and with other agencies providing services to juveniles, etc.
3. Commitments to change and innovation.
4. Staff identification with members of executive cadre.
5. Degree of centralization of policy-making and decentralization of operational decision-making.
6. Relative power of rehabilitation-treatment cadre in organizational decision-making.

Organizational structure and program technology.

1. Specificity of objectives for components within units' technological systems.
2. Staff division of labor based on differential task and skill requirements.
3. Levels of staff skills and amount of training required for technological tasks.
4. Articulation between technical system components, and between these and staff structures and roles.
5. Amount of resources allocated to rehabilitation-treatment technologies.
6. Ratio of staff performing rehabilitation-treatment roles to those performing maintenance and custodial roles.
7. Patterns of interaction and communication among sub-units within the organization -- formal and informal.
8. Degree of adaptability of the technical systems to changes in both offender populations and environmental conditions.
9. Offender truancy or "AWOL" rates.

Staff-Offender relations.

1. Ratio of staff to offenders -- both total and for staff sub-groups.
2. Degree of support by staff for development of informal offender systems.
3. Extent of positive primary group relations between staff and offenders and staff-offender social distance.
4. Degree of staffs' non-stigmatizing perceptions of offenders.
5. Offenders' perceptions of staff as helping persons.
6. Types of rewards and punishments employed by staff and organizational controls over their use.
7. Ratio of gratification to deprivation for offenders.
8. Degree of offenders' participation in decision-making about major areas of organizational activity.
9. Extent to which due process procedures are adhered to in managing daily living and control of offenders.

Inter-organizational Processes and Efforts. Our concern at this level will be directed toward the effectiveness of service units in relation to their environments. It is recognized that the nature of units' environments, including those under private auspices, may be variously manifested or defined within and between the states. Effectiveness at this level is to be assessed in terms of organizational adaptability; organizational legitimacy; relations with regulatory groups; mobilization of resources; control over input and output (including offenders); relations with complementary organizations; and total level of exchanges with other units and organizations in the external environment.

1. Degree of resource control by the service unit.
2. Degree of autonomy in determining service unit policies and program content.
3. Routinization of linking mechanisms for inter-organizational exchange.
4. Degree of monitoring of exchanges.
5. Provision for feedback and adaptation.
6. Degree of congruency in expectations between unit and external agencies relevant to the unit.
7. Stability and rate of increment in resources for priority goals.

Many of the above criteria require greater specification before they can be measured reliably and validly. It is also expected that the total number of criteri can be reduced as more information is obtained about each type of program so as to be able to determine relative criticality for comparative assessment of effectiveness

Conclusion

We have attempted to delineate some of the major issues and dilemmas in the evaluation of effectiveness in juvenile corrections. The analysis of several programs reported in the literature highlighted the serious problems in evaluation methodology, as well as the grim picture with respect to knowledge about technologies that will produce greater outcome success. To achieve this end, evaluation of process is as important as is measurement of outcome per se. Only when we can establish linkages between events within the program and subsequent outcomes can we have the knowledge that is needed for policy recommendations. Yet, most evaluative research continues to expend more resources on unrelated measurement of inputs rather than process and outputs.

Correctional organizations employ highly varied people-processing and people-changing technologies to achieve both manifest and latent ends. Yet, criteria for choice among technologies remain unclear and often are non-existent. Instead, choices are based on fads and hunches without reference to input characteristics or output objectives.

Thus far evaluation in human service organizations has focussed primarily on program implementation phases rather than on assessment of program design, planning and formulation. Weiss (1972) argues that greater priority should be given to the latter so that knowledge will be obtained about how to avoid or cope with organizational and environmental problems which often occur. The evaluation literature is filled with reports on major organizational changes which have occurred in the middle of the evaluation effort. In most cases these were unanticipated and researchers decried the problems created in adaptation, measurement, and so forth.

Such change needs to be anticipated as a likely rather than an unusual event. Evaluative research will also be different and probably more difficult than other types of social science research because it deals directly with reality in settings where research is not the primary activity.

The politics of evaluation were dealt with only peripherally in this paper, but they are of critical importance to the researcher today whether he or she likes it or not (Rossi and Williams, 1972; Weiss, 1972). The scientist must anticipate how his findings will be read, misunderstood, ignored and distorted. He must be prepared to explain at other than a superficial level why the results were obtained and what alternative meanings they may have. Over and over the findings from evaluative research report that the null hypothesis of no change was supported. The question then arises: does that mean that the organization had no impact or that the proper dimensions of organizational behavior were not measured? Only infrequently are such questions considered.

Evaluation inevitably has political implications, for it is the means by which the character of a program or organization is described and analyzed. Social values are always involved and must be addressed as such. Many recent observers have referred to the increasing politicization of juvenile justice systems in several countries. This phenomenon will further add to problems in evaluation for increasing pressures will be exerted on researchers. For example, we are in a period of rapidly escalating costs in all correctional programs, so legislators and others are looking for information which will be of use in resource allocation decisions. When the cost of institutionalization for a juvenile offender now exceeds $20,000 per year (as it does in several places), decision-makers will exert great pressure to obtain evaluative data about both process and outcome. Of even greater importance than costs are the social consequences of correctional experience for individual youth and for the society as a whole. Social scientists must be willing to deal with these and other value and policy questions. Such situations provide opportunities to enhance the development of better evaluation methodologies and the utilization of research findings.

REFERENCES

Adamek, Raymond J. and Edward Z. Dager.
1968 "Social Structure, Identification and Change in a Treatment-Oriented Institution." American Sociological Review, 33, 931-944.

American Association of University Women.
1970 Survey of Prisons for Women in the United States. Washington, D.C.: American Association of University Women.

American Friends' Service Committee.
1971 Struggle for Justice: A Report on Crime and Punishment in America. New York: Hill and Wang.

Aronson, Sidney and Clarence Sherwood.
1967 "Researcher vs. Practitioner: Problems in Social Action Research." Social Work, 12 (October):89-96.

Bailey, Walter C.
1966 "Correctional Outcome: An Evaluation of 100 Reports." Journal of Criminal Law, Criminology and Police Science, 57 (June):153-160.

Baum, Martha and S. Wheeler.
1968 "Becoming an Inmate." pp. 153-185 in Stanton Wheeler (ed.), Controlling Delinquents. New York: Wiley.

Brim, Orville and S. Wheeler.
1966 Socialization After Childhood. New York: Wiley.

Burghardt, Stephen, Rosemary C. Sarri and Carl Gohlke.
1971 "A Comparative Assessment of Probation Practices and Perspectives." (September), in A Comparative Study of Federal Correctional Programs for Young Offenders. Washington, D.C.: U.S. Bureau of Prisons (forthcoming).

California Youth Authority.
1973 "Follow-up of Wards Discharged From California Youth Authority During 1965." Sacramento: California Youth Authority Research Report, No. 64.

Campbell, Donald T. and Julian Stanley.
1966 Experimental and Quasi-Experimental Design for Research. Chicago: Rand, McNally.

Caro, Francis G. (ed.)
1971 Readings in Evaluation Research. New York: Russell Sage.

Christie, Nels.
1968 Scandanavian Studies in Criminology. London: Tavistock, Vol. 2, 73-107.

Cicourel, Aaron and John Kitsuse.
1964 The Educational Decision Makers. Indianapolis: Bobbs - Merrill.

Clemmer, Donald.
1940 The Prison Community. New York: Holt, Rinehart and Winston.

Cressey, Donald R.
1965 "Prison Organizations." pp. 1023-1067 in James Marsh (ed.), Handbook of Organizations. Chicago: Rand, McNally.

Cressey, Donald R.
1961 The Prison: Studies in Institutional Organization and Change. New York: Holt, Rinehart and Winston.

Deniston, O.L., I.M. Rosenstock and V.A. Getting.
1968 "Evaluation of Program Effectiveness." Public Health Reports, 83:4; 323-335.

Emerson, Robert.
1969 Judging Delinquents: Context and Process in Juvenile Court. Chicago: Aldine.

Empey, LaMar and Maynard T. Erickson.
1972 The Provo Experiment: Evaluating Community Control of Delinquency. Lexington, Mass.: C.C. Heath & Co.

Empey, LaMar and Steven Lubeck.
1971 The Silverlake Experiment. Chicago: Aldine.

Erickson, R.J., W.J. Crow, L.A. Zurcher, A.V. Connett and W.D. Stillwell.
1971 The Offender Looks at His Own Needs. La Jolla, California: Western Behavioral Sciences Institute.

Fogel, David.
1969 "The Fate of the Rehabilitative Ideal in California Youth Authority Decisions." Crime and Delinquency, 15 (October): 479-498

Gafni, Miriam and Barney Welsh.
1969 "Post-Conviction Problems and the Defective Delinquent." Villanova Law Review, 12 (Spring):546-602.

Glaser,Daniel.
1964 The Effectiveness of a Prison and Parole System. Indianapolis: Bobbs-Merrill.

Glaser, Daniel and Vincent O'Leary.
1966 Personal Characteristics and Parole Outcome. Washington, D.C.: Office of Juvenile Delinquency and Youth Development, U.S. Department of Health, Education and Welfare.

Gottfredson, Don.
1967 "Current Information Bases from Evaluating Correctional Programs." pp. 28-33 in Research in Correctional Rehabilitation. Washington: Joint Commission on Correctional Manpower and Training.

Guttentag, Marcia.
1971 "Models and Methods in Evaluation Research." (Unpublished paper, New York University).

Ingraham, Bastom, L. and Gerald L. Smith.
1972 "The Use of Electronics in the Observation and Control of Human Behavior and Its Possible Use in Rehabilitation and Parole." Issues in Criminology, 7 (Fall):35-54.

James, George.
1962 "Evaluation in Public Health Practice." American Journal of Public Health, (July), pp. 1145-1154.

Jesness, Carl F., William DeRisi, Paul M. McCormick and Robert F. Wedge.
1972 The Youth Center Research Project. Sacramento: American Justice Institute in cooperation with California Youth Authority, (July).

Jones, Wyatt C. and Edgar F. Borgotta.
1972 "Methodology of Evaluation." in F.J. Mullin and J.R. Dumpson, et. al (eds.), Evaluation of Social Intervention. San Francisco: Jossey, Bass.

Kandel, Denise Hystryn and Richard Hays Williams.
1964 Psychiatric Rehabilitation: Some Problems of Research. New York: Atherton Press.

Kassebaum, Gene, David Ward, and D. Wilner.
1970 Prison Treatment and Its Outcome. New York: Wiley.

Kitano, Harry H.L.
1963 "The Concept of 'Precipitant' in Evaluative Research." Social Work, VIII (October):34-38.

Kittrie, N.
1971 The Right to Be Different: Deviance and Enforced Therapy. Baltimore: Johns Hopkins University Press.

Krasner, Leonard.
1962 "Behavior Control and Social Responsibility." American Psychologis 17:4, 199-203.

Lerman, Paul.
1968 "Evaluative Studies of Institutions for Delinquents: Implications for Research and Social Policy." Social Work, 13 (July):55-64.

Lohman, Joseph D., Albert Wahl and Robert M. Carter.
1967 The San Francisco Project. Berkeley, California: University of California School of Criminology.

Mandell, Nathan.
1965 "Recidivism Studied and Defined." Journal of Criminal Law, Criminology and Police Science, (March):59-66.

Meyer, Henry J. Edgar F. Borgatta and Wyatt C. Jones.
1965 Girls at Vocational High: An Experiment in Social Work Intervention New York: Russell Sage Foundation.

Miller, Walter B.
1962 "The Impact of a Total-Community Delinquency Control Project." Social Problems, 10 (Fall):168-191.

Minnesota Department of Corrections.
1972 "A Follow-up Study of Boys Participating in the Positive Peer Culture Program at the Minnesota State Training School for Boys: An Analysis of 242 Boys Released During 1969." St. Paul: Minnesota Department of Corrections.

Moberg, David O. and Richard Ericson.
1972 "A New Recidivism Outcome Index." Federal Probation, XXXVI (June): 50-56.

Moos, Rudolf H.
1968 "Assessment of Social Climates of Correctional Institutions." Journal of Research in Crime and Delinquency, 5, 174-188.

Morris, Albert.
1971 A Correctional Administrators' Guide to the Evaluation of Correctional Programs, Bulletin No. 21. Boston, Mass.: Massachuse Corrections Association, 33 Mt. Vernon Street.

Morris, Narval and Gordon Hawkins.
1969 The Honest Politician's Guide to Crime Control. Chicago: Universi of Chicago Press.

Mott, Paul.
1972 The Characteristics of Effective Organizations. New York: Harper and Row.

National Conference on Criminal Justice.
1973 Working Papers for the National Conference on Criminal Justice. Washington, D.C.: LEAA.

Pappenfort, Donald and Dee Kilpatrick.
1970 A Census of Residential Institutions in the United States, Puerto Rico and the Virgin Islands, 1966. Chicago: University of Chicago School of Social Service Administration.

Riecken, Henry W.
1972 "Memorandum on Program Evaluation." in Evaluating Action Programs: Readings in Social Action and Education, edited by Carol H. Weiss, Boston: Allyn and Bacon, Inc. pp. 85-105.

Robinson, James and Gerald Smith.
1971 "The Effectiveness of Correctional Programs." Crime and Delinquen 17, 67-80.

Rossi, Peter H. and Walter Williams.
1972 Evaluating Social Programs: Theory, Practice and Politics. New York: Seminar Press.

Sarri, Rosemary C., John Tropman, Matthew Silberman, Edward J. Pawlak and Ken Badal.
1970 Client Careers and Public Welfare Structures. A Progress Report, U.S. Department of Health, Education, and Welfare, Social and Rehabilitation Service, Grant No. CRD-425-C1-9. Ann Arbor, Michigan: School of Social Work

Sarri, Rosemary C. and Robert D. Vinter.
1967 "Organizational Requisites for a Socio-Behavioral Technology," pp. 87-100 in E.J. Thomas (ed.), The Socio-Behavioral Approach and Applications to Social Work. New York: Council on Social Work Education, 1967.

Sarri, Rosemary C. and Robert D. Vinter.
1972 National Assessment of Juvenile Corrections: Research Design Statement. Ann Arbor: University of Michigan.

Schafer, Walter and Carol Alexa.
1971 Tracking and Opportunity. San Francisco: Chandler Press.

Schrag, Clarence.
1971 Crime and Justice: American Style. Rockville, Md.: National Institute of Mental Health, Center for Studies of Crime and Delinquency, Publication No. HSM-72-9052.

Schwitzgebel, Ralph.
1967 "Issues in the Use of an Electronic Rehabilitation System with Chronic Residents." (Unpublished paper)

Seidman, David and Michael Couzens.
1972 "Crime, Crime Statistics and the Great American Anti-Crime Crusade: Police Misreports of Crime and Political Pressures." Paper presented at the 1972 Annual Meeting of the American Political Science Association. Copyright 1972 by American Political Science Association.

Shapiro, Michael.
1972 "The Use of Behavior Control Technologies: A Response." Issues in Criminology, 7 (Fall):55-93.

Sherwood, C.C.
1967 "The Testability of Correctional Goals." in Research in Correctional Rehabilitation. Washington, D.C.: Joint Commission on Correctional Manpower and Training, pp. 24-32.

Slosar, John.
1972 "Prisonization: Social Relationships and Leadership Patterns in Two Federal Youth Centers." (Unpublished paper).

Southern California Law School.
1972 "Conditioning and Other Technologies Used to 'Treat?' 'Rehabilitate' 'Demolish?' Prisoners and Mental Patients." Notes: Southern California Law Review, 45, 616-685.

Street, David, Robert D. Vinter and Charles Perrow.
1966 Organization for Treatment. New York: Free Press of Macmillan.

Suchman, Edward.
1967 Evaluation Research. New York: Russell Sage Foundation.

Terry, Robert M.
1970 "Discrimination in the Handling of Juvenile Offenders by Social Control Agencies." pp. 79-92 in Peter G. Garabedian and Don C. Gibbons (eds.), Becoming Delinquent: Youthful Offenders and the Correctional System. Chicago: Aldine.

Thomas, Charles and Samuel C. Foster.
1972 "Prisonization in the Inmate Contra-culture." Social Problems, 20 (Fall):229-239.

Ullmann, Leonard P.
1967 Institution and Outcome: A Comparative Study of Psychiatric Hospitals. London: Pergamon Press.

U.S. Bureau of the Census.
1970 General Population Characteristics: Final Report. PC(1)-B40.

Vinter, Robert D.
1963 "Analysis of Treatment Organizations." Social Work, VIII (July):3

Warren, Marguerite, et. al.
1966 Community Treatment Project: An Evaluation of Community Treatment for Delinquents. Sacramento, California: California Youth Authority CTP Research Report No. VII.

Ward, David and Gene Kassebaum.
1972 "On Biting the Hand That Feeds: Some Implications of Sociological Evaluations of Correctional Effectiveness." in C. Weiss (ed.), Evaluating Action Programs. Boston: Allyn, Bacon, Inc.

Webb, Gene E.
1972 "Re-thinking Macro-system Intervention." in Edward J. Mullen, James R. Dumpson, et. al (eds.), Evaluation of Social Intervention San Francisco: Jossey-Bass.

Weiner, Norman L. and Charles V. Willie.
1971 "Decisions by Juvenile Officers." American Journal of Sociology, 77 (September):199-210.

Weiss, Carol H.
1972 Evaluating Action Programs: Readings in Social Action and Education. Boston: Allyn and Bacon, Inc.

Wolfgang, Marvin.
1972 "Making the Criminal Justice System Accountable." Crime and Delinquency, 18 (January):15-22.

Wolins, Martin.
1960 "Measuring the Effect of Social Work Intervention." pp. 247-273 in N. Polansky (ed.) Social Work Research. Chicago: University of Chicago Press.